基于图像深度学习的无线电信号识别

周鑫※　　何晓新　　郑昌文

（中国科学院软件研究所 天基综合信息系统重点实验室　北京　100190）

摘　要：本文创新性地提出了一种利用图像深度学习解决无线电信号识别问题的技术思路：首先把无线电信号具象化为一张二维图片，将无线电信号识别问题转化为图像识别领域的目标检测问题；进而充分利用人工智能在图像识别领域的先进成果，提高无线电信号识别的智能化水平和复杂电磁环境下的识别能力。基于该思路，本文提出了一种基于图像深度学习的无线电信号识别算法——RadioImageDet 算法。试验结果表明，该算法能有效识别无线电信号的波形类型和时/频坐标，在实地采集的 12 种、4740 个样本的数据集中，识别准确度达到 86.03%，mAP 值达到 77.72，检测时间在中等配置的桌面计算机上仅需 33 毫秒，充分验证了思路的可行性和算法的有效性。

关键词：无线电信号识别；深度学习；射频机器学习；卷积神经网络；图像目标检测

中图分类号：

文献标识码：A

DOI：

Radio Signal Recognition Based on Image Deep Learning

ZHOU Xin※　　HE Xiaoxin　　ZHENG Changwen

(*Science & Technology on Integrated Information System Laboratory,*

Institute of Software Chinese Academy of Sciences, Beijing 100190, China)

Abstract: This paper innovatively proposes a technical idea that uses image deep learning to solve the problem of radio signal recognition: first, it transforms the radio signal into a two-dimensional picture, and transforms the radio signal recognition problem into the object detection problem in the field of image recognition; then, it makes use of the advanced achievements about image recognition to improve the intelligence and ability of radio signal recognition in complex electromagnetic environment. Based on this idea, a novel radio signal recognition algorithm named *RadioImageDet* is proposed in this paper. The experimental results show that the algorithm can effectively identify the waveform types and time/frequency coordinates of radio signals. After training and testing on the self-collected data set with 12 types and 4740 samples, the accuracy reaches 86.03% and the mAP value reaches 77.72, while the detection time is only 33 milliseconds on the medium configured desktop computer.

Key words: Radio Signal Recognition; Deep Learning; Radio Frequency Machine Leaning (RFML); Convolutional Neural Network (CNN); Image Object Detection

1 引言

近年来，随着无线通信技术的发展，特别是物联网的兴起，无线电波已经成为连通万物的重要载体。然而，无线电波开放性的特点，使其容易受到干扰和非法利用，导致正常通信系统受扰甚至中断，威胁社会和国家安全。因此，加强无线电监测与识别，特别是在机场、边境等重点区域以及重大活动现场的无线电监管，具有重要的现实意义和迫切需求。

无线电信号识别技术最初主要用于军事电子战和政府频谱监管领域，随着近年来频谱共享需求的提升和认知无线电技术的发展，又逐渐成为商业通信领域关注的焦点[1]。早期的无线电信号识别，是利用各种频谱采集设备采集空中射频信号，经过处理后，以频谱图、瀑布图、余晖图等可视化方式展现出来，然后由专业人员分析信号时频特征并寻找目标信号[2]。这种做法对操作人员的专业素质要求非常高，而且当监测时间变长或无线电信号较多时，人工分析效率和准确率也会大幅下降。

本研究由国防科技创新特区基金资助（编号：17-163-11-ZT-003-019-01）

※通信作者：周鑫　email 地址：zhouxin@iscas.ac.cn

目前比较流行的无线电信号识别方式，是先由本领域专家精心挑选信号关键特征，如功率谱密度最大值[3]、信号包络峭度[4]、瞬时相位标准差[5]、相位脉冲个数[4]、循环谱[6-7]、高阶矩[8-9]等时/频/谱域信号特征，然后利用传统的信号处理算法计算特征参数值，最后再基于固定规则或者机器学习的方法进行分类[10]。通常，一个模型需要选取其中的几项甚至十几项特征作为模型输入，计算复杂度非常高，很难实际部署使用。同时，受信号波形多样性和多径衰落效应的影响，该方法在特征选择和判决准则方面也缺乏普适性。

众所周知，人工智能在模式识别任务中具有独特优势，近年来在图像/语音识别、自然语言处理、医疗、金融等很多领域取得了令人震惊的成就，引起了社会各界的热切关注。无线电信号识别，本质上也是一种特殊的"模式识别"，因此我们认为，可以将人工智能与传统无线电信号识别技术深度融合，利用人工智能的机器学习方法，特别是深度学习算法，自动提取无线电波的模式特征，避免基于经验的人工特征提取环节，提高复杂电磁环境下无线电信号的识别能力。

基于这一设想，本研究创新性地提出一种利用图像深度学习方法解决无线电信号识别问题的技术思路：首先把无线电信号具象化为一张二维图片，将无线电信号识别问题转化为图像识别领域的目标检测问题；进而充分利用人工智能在图像识别领域的先进成果，提高无线电信号识别的智能化水平和复杂电磁环境下的识别能力。基于该思路，本文提出了一种基于图像深度学习的无线电信号识别算法——RadioImageDet 算法，并通过实物系统试验，验证了思路的可行性和算法的有效性。

2 相关研究工作

早在上世纪九十年代，传统机器学习算法，如朴素贝叶斯、神经网络、决策树等，就已经在无线电调制识别领域中得到了应用[12-14]。但受限于当时的技术发展水平，仅仅是利用机器学习算法，根据人工提取出的信号特征进行分类，具体如图 1 所示。

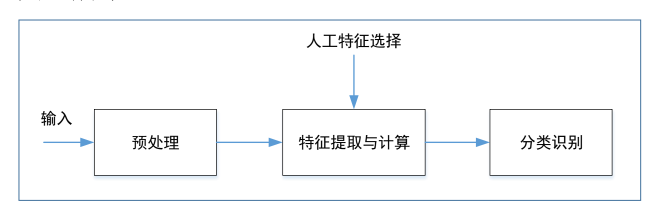


图 1 传统基于机器学习的无线电信号识别流程

近两年，随着计算机技术和深度神经网络技术的飞速发展，这一现状被逐渐打破。2016 年，受到深度神经网络在图像识别领域巨大成功的鼓舞，O'shea 及其团队连续发表多篇论文，研究深度学习技术在无线电通信与识别领域应用的方法与问题[15-17]，开启了无线电识别从经验驱动的人造特征范式到数据驱动的表示学习范式的新纪元。同年，美国国防先进研究项目局（DARPA）宣布，将于 2017 年启动 SC2 挑战赛（The Spectrum Collaboration Challenge，频谱协作挑战），旨在"将先进的机器学习能力应用于无线电领域，优化频谱使用策略，以缓解电磁频谱拥挤状态"[18]。这是本世纪 DARPA 组织的第三次挑战赛，前两次分别是 21 世纪初的自动化月球探测器挑战赛，以及 2015 年的机器人挑战赛。

2017 年，DARPA 又频频出手，推动 AI（Artificial Intelligence）与无线电的融合。2017 年 3 月，DARPA 借助认知无线电领域的著名学术会议——DySPAN 会议，组织了一场"ModRec"竞赛，要求各挑战者识别会议现场随机生成的 20 多种不同调制方式的信号波形，并声称"这是以更高效的方式管理电磁频谱的第一步"。8 月，DARPA 又启动了"射频机器学习系统"（RFMLS）项目。作为 SC2 项目的补充，RFMLS 项目旨在"建立一种射频鉴别能力，能够从信号嘈杂聚集的频谱中分辨出独特和特殊的信号"[19]，未来将用于民用和军事领域。

下面从三个方面介绍与本研究相关的研究工作。

2.1 基于 CNN 的调制方式识别

O'shea 在文献[15]中，提出了一种基于端到端 CNN 的调试方式识别模型。该模型以时域 I/Q 采样为输入，经过两层卷积层和两层全链接层处理，输出信号的调制方式，具体结构如表 1 所示。论文利用 GnuRadio

仿真生成了 11 种不同调制方式的信号，作为训练和测试数据。通过与朴素贝叶斯、K 近邻、决策树、多层感知机、支持向量机等传统机器学习方法进行对比，发现新算法与传统算法的最佳性能相当（传统算法采用以循环矩为基础构造的 32 个特征作为算法输入），证明了端到端 CNN 用于自动调制识别的有效性。

表 1 O'Shea 模型结构

网络层类型	输入尺寸	参数	激活函数
Conv 层	128×2	3×1 卷积核 256 个特征映射 Dropout 50%	ReLU
Conv 层	256×126×2	3×2 卷积核 80 个特征映射 Dropout 50%	ReLU
Dense 层	9920×1	256 个神经元 Dropout 50%	ReLU
Dense 层	256×1	11 个神经元	softmax

在此基础上，文献[20]采用分层识别的方法进行改进，即先用一个 CNN 网络区分模拟调制和数字调制，然后在该先验知识基础上，再用一个 CNN 网络识别具体的调制类型，最后再用一个 CNN 网络识别调制的阶数。文献[21-22]从网络结构的角度进行改进，在 CNN 网络基础上，增加了时序特征的考虑，采用 CNN+LSTM 的网络结构进行调制方式识别。文献[23]在 O'shea 模型的基础上，在全连接层之前增加了一些边缘信息（即传统的循环平稳特征参数）作为输入，用于辅助 CNN 识别。该方法摒弃了完全以原始信息作为模型输入的前提条件，将传统基于机器学习的自动调试识别方法与新型的 CNN 网络结合，虽然提高了识别准确率，但遗留了很强的专家系统印记，且计算复杂度较高。

2.2 无线电个体识别

文献[15, 20-23]均属于调制方式识别（Modulation Recognition），其输出需要再次联合其它信息，才能形成对电子侦察、频谱监管等应用有直接作用的结果；而文献[24-27]则属于更高一层的无线电个体识别（Radio Individual Identification），其结果可直接用于上述应用场景，因此更加符合"端到端"识别的理念。

文献[24]利用信号矩阵的次要特征值区分信号，通过矩阵特征值求解法进行无线电信号识别。文献[25]利用物理波形和射频器件的非线性杂散发射特征区分信号发射源，并用相位空间重构法（RPS）进行分析识别。文献[26]利用信号功率谱密度（PSD）的相似性区分信号，并采用基于神经模糊（neuro-fuzzy）的相似度测量方法进行分析识别。以上文献均属于传统的基于专家系统的无线电识别方法，而 Schmidt 在文献[27]中提出了一种基于 CNN 的无线电通信标准识别模型，该模型结构与 O'Shea 模型相似，但却以频域快速傅里叶变换（FFT）变换结果为输入，直接输出无线电通信标准名称，原文中涉及 802.11、802.15.1、802.15.4 三种通信标准。文献[28]也采用类似的 CNN 网络，但是以 Choi-Williams 时频分布（CWD）变换结果作为模型输入，用于识别线性调频（LFM）、Frank 序列、Costas 码等 8 种常见雷达波形。本研究也属于无线电个体识别，其输入为包含时/频信息的频谱瀑布图，输出为信号的波形名称及其时/频坐标。

2.3 基于深度学习的图像目标检测

图像目标检测的任务就是找出图像中所有感兴趣的物体，并确定他们的位置和大小，是计算机视觉领域的核心问题之一。2013 年以前，以 DPM（多尺度形变部件模型）为代表的传统目标检测方法一直处于强势地位。2014 年，R. Girshick 提出了 R-CNN（Region proposal CNN）算法框架[29]，使得目标检测的准确率取得巨大突破，并开启了基于深度学习的目标检测的研究热潮。随后，Fast R-CNN[30], Faster R-CNN[31], FPN[32], YOLOv2[33], SSD[34]等一系列研究成果纷纷涌现，在检测速度、检测准确率等方面均有较大提高。目前，基于深度学习的图像目标检测主要分为两类：一类是以 R-CNN、Faster R-CNN 为代表的基于区域提

取（region proposal）的目标检测算法，另一类是以 YOLOv2、SSD 为代表的基于回归问题的端到端目标检测算法。由于前者需要把检测过程分为区域提取和检测分类两步，而后者直接在一个卷积网络中实现目标的定位和识别，简化了网络结构，在保证准确率的前提下，检测速度达到了实时检测任务的要求，因此，本研究将以 YOLOv2 为基础，设计无线电信号识别模型。

3 RadioImageDet 算法模型设计

本文提出了一种基于图像深度学习的无线电信号识别算法——RadioImageDet 算法。该算法基于一种端到端卷积神经网络，能够在时/频混叠的复杂电磁环境下，同时识别出环境中存在的多个信号的波形类型，并可定位出各自的起止时间和频率范围，其算法结构如图 2 所示。

RadioImageDet 算法由预处理模块和端到端目标检测模块两部分组成。预处理模块以观察窗口的原始 I/Q 采样数据为输入，经过离散傅里叶变换（DFT），将时域数据转变成更易识别的频域表达；再经过数据图像化处理，将连续多次 DFT 的结果组合成一张二维频谱瀑布图。端到端目标检测模块以预处理所得的频谱瀑布图为输入，经过 CNN 特征提取层，获取具象化的信号特征；再经过目标检测层，将抽象的特征转化为信号波形类型、时/频坐标等有确定含义的输出结果。

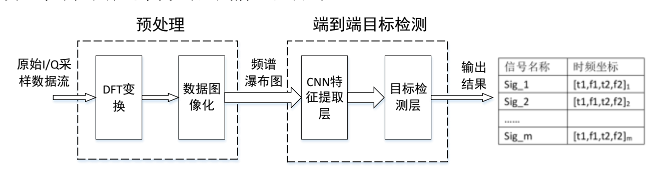


图 2 RadioImageDet 算法结构图

3.1 预处理模块

RadioImageDet 算法运行，首先要进行数据预处理。预处理的目的，是为了将原始 I/Q 采样数据变换为一种更有利于自动特征提取和识别的表达形式。为了充分发挥卷积神经网络在特征提取方面的优势，并降低计算复杂度，我们希望预处理过程是无损且容易计算的。根据离散信号处理理论和现代通信系统原理，原始的 I/Q 采样数据就是数字通信系统后端处理的基本输入，任何复杂的数字信号处理算法都是在此基础上经过数学变换完成，因此它天然蕴含了无损信息。这也是 O’shea 等人直接选用 I/Q 采样数据作为神经网络输入的原因[15, 20]。从数学上讲，离散傅里叶变换（DFT）是对有限长 I/Q 采样数据的无损变换，如公式（1）所示，同时 DFT 变换的结果作为信号的频域表达形式，更有利于以电磁辐射为基础的无线电信号的识别[35]。因此，本文首先对原始 I/Q 采样数据进行 DFT 变换，得到信号的频域表达，然后再进行后续的处理，这与文献[26, 27]的思路一致。

$$\begin{aligned} DFT:\ & X[k]=\sum_{n=0}^{N-1}x[n]e^{-jk(2\pi/N)n},\ k=0,1,\ldots,N-1 \\ IDFT:\ & x[n]=\frac{1}{N}\sum_{k=0}^{N-1}X[k]e^{jk(2\pi/N)n},\ n=0,1,\ldots,N-1 \end{aligned} \tag{1}$$

一次 DFT 运算的结果是一个一维复数序列，可以形象地认为是“一条线”，它表示某一时间点的频谱密度谱。但是，无线电信号波形总是随时间变化，仅观察一个“时间点”的特征很难准确识别信号，因此将连续多个时间点的观测结果进行堆叠组合，变成“一个面”，就可以同时表达时域和频域的信息特征，有助于提高信号识别率。在这个二维平面中，横轴表示频率，纵轴表示时间，元素值就是 DFT 变换的结果值。如果将元素值按图像灰度值进行映射，就可以绘出一张可视化的图片，该图片就是频谱瀑布图，具体如图 3 所示。因此，本文在 DFT 变换之后，增加了一个数据图像化处理过程，将连续多组 DFT 变换的结果组合成频谱瀑布图的格式，增加单个样本的时域信息量，便于后续卷积神经网络的处理和信号识别。

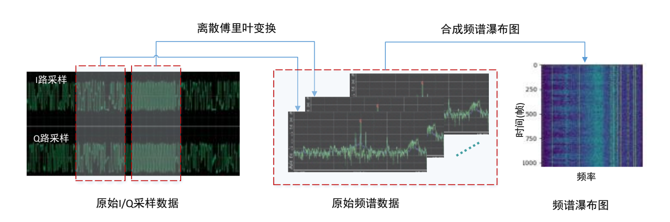


图 3 频谱瀑布图合成示意图

理论上讲，频谱瀑布图的尺寸为 M×N×2。其中，2 表示图片为两个通道，分别为 I 通道和 Q 通道；N 表示离散傅里叶变换的点数；M 表示一张频谱瀑布图内合并的离散傅里叶变换结果的组数，其大小反映了一个样本的时间跨度，M 值越大，表征的时间跨度越长，计算也越复杂。为了方便运算，本文作以下三个简化操作：

1）为了简化图像转化操作，直接将 I/Q 两通道按公式（2）合并为一个通道，合并后的物理含义为功率密度（PSD）；

$$S=\sqrt{I^2+Q^2} \tag{2}$$

2）为了利用快速傅里叶变换（FFT）算法加快运算速度，N 值需取 2 的正整数次幂，本文中取 N=512；

3）为了匹配某些基于 CNN 的图像目标检测模型，取 M = N，构成一个正方形图片。

因此，简化后的频谱瀑布图尺寸为 N×N，其中 N 为 2 的正整数次幂。

3.2 端到端目标检测模块

经过预处理，抽象的无线电信号被具象化为一张形象的二维图片，接下来就要对该图片进行特征提取，识别出图片中存在的信号及其位置。这项工作的本质与计算机视觉中的目标检测任务完全相同，所不同的是，该图片是无线电信号经过 DFT 变换后合成的频谱瀑布图，而不是自然界中的图片。由于输入图片的特殊性，或许存在更加合适的神经网络结构进行识别和目标检测，但本文暂时不对其进行深入研究，仅以目前比较流行的 YOLOv2 模型[33]为基础，通过简单参数修改，构造无线电信号识别网络模型，称为 RadioYOLO 模型。

如图 4 所示，RadioYOLO 模型结构主要参考了 Inceptionv1 网络，仅使用 3×3 和 1×1 卷积核尽力压缩模型参数量，并在每次池化操作后将特征通道数翻倍，以尽量保持特征的完整性。同时，借鉴 Inceptionv2 的思想，对每一个卷积层的输出都使用批归一化（Bath-Normalization）处理，增加了模型的鲁棒性和训练速度，并可代替 dropout 防止过拟合；借鉴 ResNet 的跳跃连接的思想，将 conv13 的输出整形后，直接与 conv20 进行通道合并，共同输出给下一层，使得输出层能够保留更高分辨率的特征，并降低“梯度消失”概率；使用 leakyReLU 非线性激活函数，减少训练中神经元死亡的概率。

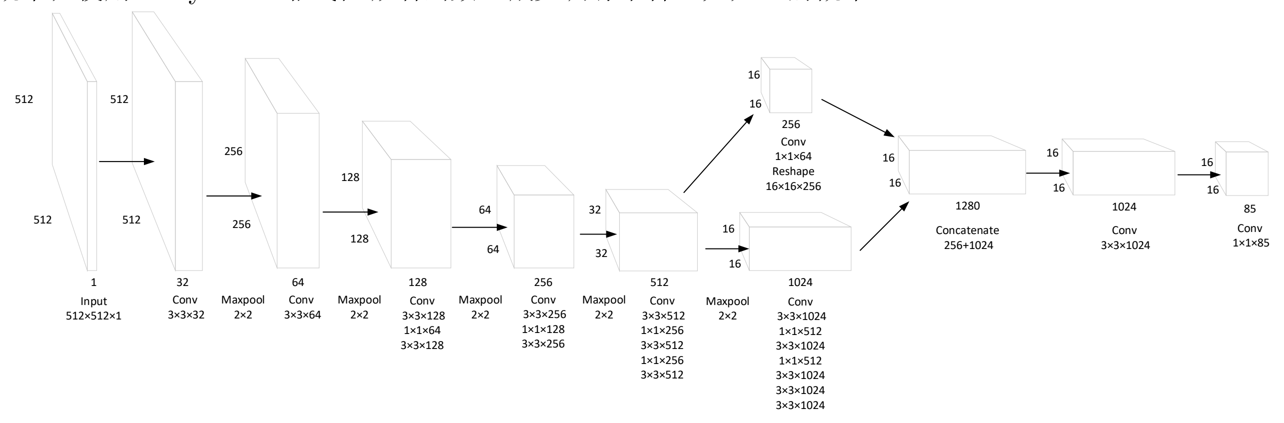


图 4 RadioYOLO 模型结构

与 YOLOv2 模型相比，RadioYOLO 的主要改动包括以下几点：

1） 模型输入的图片尺寸：根据 FFT 变换的尺寸，将输入尺寸改为 512×512；

2） 归一化锚框的尺寸：根据无线电信号的公共特征，选择 5 个锚框（width, height），分别是：

a) （0.015， 0.9）：匹配窄带持续信号

b) （0.5， 0.9）：匹配宽带持续信号

c) （0.025， 0.05）：匹配窄带瞬时信号

d) （0.469， 0.01）：匹配宽带瞬时信号

e) （0.469， 0.1）：匹配宽带瞬时信号

3） 模型输出的张量尺寸：根据输入尺寸和数据集特征，推导出新的输出尺寸为 16×16×85。

表 2 RadioYOLO 模型结构

（每个卷积层均作批归一化处理和 leakyReLU 激活，尺寸值省略通道数）

层名称	输入尺寸	尺寸/步进	卷积核数
Conv1	512×512	3×3 / 1×1	32
Maxpool1	512×512	2×2 / 2×2	/
Conv2	256×256	3×3 / 1×1	64
Maxpool2	256×256	2×2 / 2×2	/
Conv3	128×128	3×3 / 1×1	128
Conv4	128×128	1×1 / 1×1	64
Conv5	128×128	3×3 / 1×1	128
Maxpool3	128×128	2×2 / 2×2	/
Conv6	64×64	3×3 / 1×1	256
Conv7	64×64	1×1 / 1×1	128
Conv8	64×64	3×3 / 1×1	256
Maxpool4	64×64	2×2 / 2×2	/
Conv9	32×32	3×3 / 1×1	512
Conv10	32×32	1×1 / 1×1	256
Conv11	32×32	3×3 / 1×1	512
Conv12	32×32	1×1 / 1×1	256
Conv13	32×32	3×3 / 1×1	512
Maxpool5	32×32	2×2 / 2×2	/
Conv14	16×16	3×3 / 1×1	1024
Conv15	16×16	1×1 / 1×1	512
Conv16	16×16	3×3 / 1×1	1024
Conv17	16×16	1×1 / 1×1	512
Conv18	16×16	3×3 / 1×1	1024
Conv19	16×16	3×3 / 1×1	1024
Conv20	16×16	3×3 / 1×1	1024
Conv21	32×32 (以 conv13 为输入)	1×1 / 1×1	64
Reshape1	32×32	space_to_depth 2×2	256
Concatenated1	16×16 (连接 conv20 和 reshape1)	/	1280
Conv22	16×16	3×3 / 1×1	1024
Conv23	16×16	1×1 / 1×1	85

如表 2 所示，RadioYOLO 包含 23 层卷积层和 5 层池化层，输入为 512×512×1 的频谱瀑布图，输出为 16×16×K 的张量，K 的取值如公式（3）所示。

$$K = anchor_num \times (class_num + 5) \tag{3}$$

其中，$anchor_num$ 表示锚框（anchor）的个数，$class_num$ 表示待识别信号类型的个数。在本文的验证试验中，$anchor_num=5$，$class_num=12$，因此 $K=85$。

括号内的 5 表示除了信号类型外，还有 5 个维度的特征参数预测值[t_x, t_y, t_w, t_h, t_o]，分别体现信号边框的左上点坐标(b_x, b_y)、边框尺寸(b_w, b_h)和置信度($conf$)。为了提高模型训练的稳定性，t_x 等预测值不是真实值，而是一种归一化的中间值，它们与真实值的关系满足公式（4）。因此，在利用神经网络模型得到预测特征张量后，仍需通过公式（4）计算得到最终的真实预测值，完成目标检测。

$$\begin{cases} b_x = Sigmoid(t_x) + c_x \\ b_y = Sigmoid(t_y) + c_y \\ b_w = anchor_w * e^{t_w} \\ b_h = anchor_h * e^{t_h} \\ conf = Sigmoid(t_o) \end{cases} ,其中 Sigmoid(x) = \frac{1}{1+e^{-x}} \tag{4}$$

4 数据集构建

4.1 数据采集

在以往的无线电识别研究文献中，无线电数据通常来自于仿真系统，如 MATLAB 仿真软件、GnuRadio 仿真软件、信号发生器等，在实验室环境下，通过人工代码的控制，生成单一的或混合的有噪信号[15-17, 20-23, 27-28]。这种方式降低了数据获取和预处理的难度，方便利用先验信息进行数据标注，进而完成算法的验证。但是，受到信道衰落、时钟抖动、未知辐射源、接收机射频特性等复杂因素的影响，仿真数据与真实环境下的电磁环境往往存在较大差异，影响了算法的真实性和实用性。

为此，本研究利用 Gnuradio+USRP 软件无线电平台，开发了一套高速频谱采集系统，用于采集真实环境下的频谱数据，如图 5 所示。USRP 设备是由美国 NI 公司设计并生产的一款开源软件无线电模块，采用 AD9361 射频集成芯片方案，工作频段可覆盖 70MHz～6GHz。Gnuradio 是一个开源软件，利用普通计算机实现调制/解调、编解码等数字信号处理，与 USRP 配套使用，可快速构建无线通信原理验证系统。该系统通过优化数据存储算法，最高可支持 50MHz 瞬时带宽信号的分钟级无损保存，并可在预设频率集上自动扫频存储。

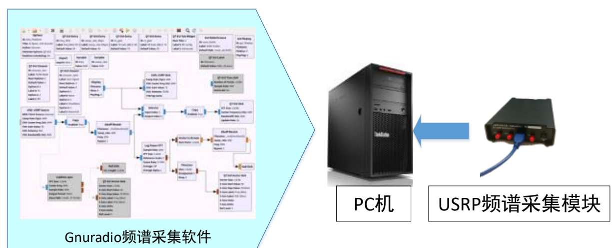


图 5 高速频谱采集系统

本文的所有训练数据和测试数据均来自该平台的实地频谱采样。采样速率为 40Msps，单频点采样时长约 5 秒，采样地点包括北京中关村地区、天津滨海新区、天津渤海湾三地共 7 个点位，关注信号为中国大陆常见 2G/3G/4G 蜂窝下行信号 12 种，在包含关注信号的频段内，每偏移 10MHz 设立一个采样频点，共计 34 个频点，数据集总大小约为 650GB。

表 3 是北京中关村地区三大运营商 2G/3G/4G 蜂窝下行频率分配表，该表是基于相关标准文件，并利用仪器实际测量观察后校对所得。结合北京、天津等地的实测结果，发现该表在中国大陆境内并不通用，各

运营商有可能在不同区域按照入网用户的数量和流量特征，在一定条件下自主调整频率资源的分配，使得不同区域、相同频点处的信号类型可能有所不同。这对本研究的核心算法没有任何影响，但会给数据标注工作带来一定麻烦。

表 3 北京地区三大运营商 2G/3G/4G 蜂窝下行频率分配表

频率范围（MHz）	信号名称
870-885	电信 2G_CDMA_下行
935-955	移动 2G_GSM900_下行
955-960	联通 2G_GSM900_下行
1805-1820	移动 2G_GSM1800_下行
1840-1860	联通 4G_FDD_下行
1860-1875	电信 4G_FDD_下行
1885-1905	移动 4G_TD-LTE
1905-1915	移动 4G_TD-LTE
2010-2025	移动 3G_TD-SCDMA
2110-2130	电信 4G_FDD_下行
2130-2135	联通 3G_WCDMA_下行
2135-2140	联通 3G_WCDMA_下行
2140-2155	联通 4G_FDD_下行
2575-2595	移动 4G_TD-LTE
2595-2615	移动 4G_TD-LTE
2615-2635	移动 4G_TD-LTE

4.2 数据预处理与标注

原始数据的格式为时域 I/Q 采样，必须先转换为频谱瀑布图，才能被 RadioYOLO 模型识别和处理。本研究利用 Python 语言和相关第三方库，按照 3.1 节的预处理流程，编写了一系列脚本代码，将采集到的 650GB 时域 I/Q 采样数据，转变为 4740 张 512×512 的灰白图片，其中训练集 3792 张，测试集 948 张。

之后，需要对训练和测试样本进行标注。数据标注本是一项繁琐的人力劳动，但是我们基于本数据集内待识别信号具有固定工作频段且持续存在的特点，设计了一种半自动的数据标注方法，大大降低了重复性工作。该方法先在每个点位、每个频点上挑选 2 张频谱图进行人工标注，然后对该点位、该频点上剩余的频谱图进行软件自动化标注，其标注坐标在两张人工标注的坐标范围内随机抖动。标注示例如图 6 所示，在以 2140MHz 为中心的 40MHz 范围内，共存在 1 个电信 4G_FDD_下行信号（部分可见）、2 个联通 3G_WCDMA_下行信号和 1 个联通 4G_FDD_下行信号，通过标注可视化边框，软件按一定格式自动在后台生成该图片的文本标注信息。

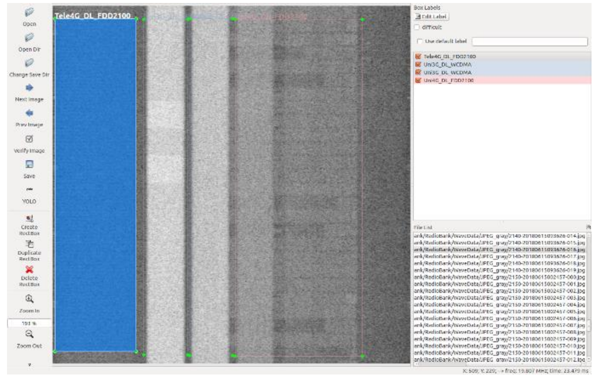


图 6 数据标注示例（2140MHz 频点）

值得一提的是，虽然本研究中数据的采集和标注都基于频率，而对于 2G/3G/4G 蜂窝等授权通信系统，

直接利用频率信息就几乎可以完成信号识别，但这并不具备通用性，且无法完成核心技术的验证。因此，我们在模型训练和检测时，刻意回避了中心频率信息，只对接收机通带内的数字基带信号进行处理。这也意味着，即使某信号被调谐到未知的非授权频段内，也不会影响本算法的检测结果。

5 结果与分析

为了验证算法性能，我们在 Tensorflow 框架下，利用 Python 语言实现了 RadioYOLO 模型，以及整个 RadioImageDet 算法，并进行模型训练与测试。数据集来自预处理后的真实频谱数据，共包含 12 种信号，其中训练样本 3792 个，测试样本 948 个，具体数据特征详见第 3 章。模型训练时，我们对训练集的前 80%进行训练，后 20%用于训练过程中的模型验证。模型训练优化器采用 adam 算法，轮次（epoch）上限设为 50，批大小（batch_size）设为 8。从实际运行结果来看，在采用 NVIDA GTX1060T 显卡加速的条件下，训练速度约为 600 毫秒/batch，每轮耗时约为 230 秒。单个样本的测试速度约为 33 毫秒，检测帧速为 30 帧/秒，基本满足实时应用需求。

识别结果的输出如图 7 所示。模型软件会根据识别的结果，自动在原图上进行标框，并注明信号类型和识别置信度。由图可见，某些信号在频谱瀑布图上的特征非常明显，通过人眼可以轻易判决，这是我们开展本研究的基本前提和动机。从检测结果上看，人眼能够轻易识别的信号，在绝大多数情况下该算法模型也能正确识别，但也有一些例外，如图 7-(g)。而人眼不易识别的信号，算法模型出错的概率更高，如图 7-(b)，但某些情况下也能正确识别，如图 7-(e)。这些现象都直观证明了利用图像深度学习方法进行无线电信号识别的可行性和有效性！

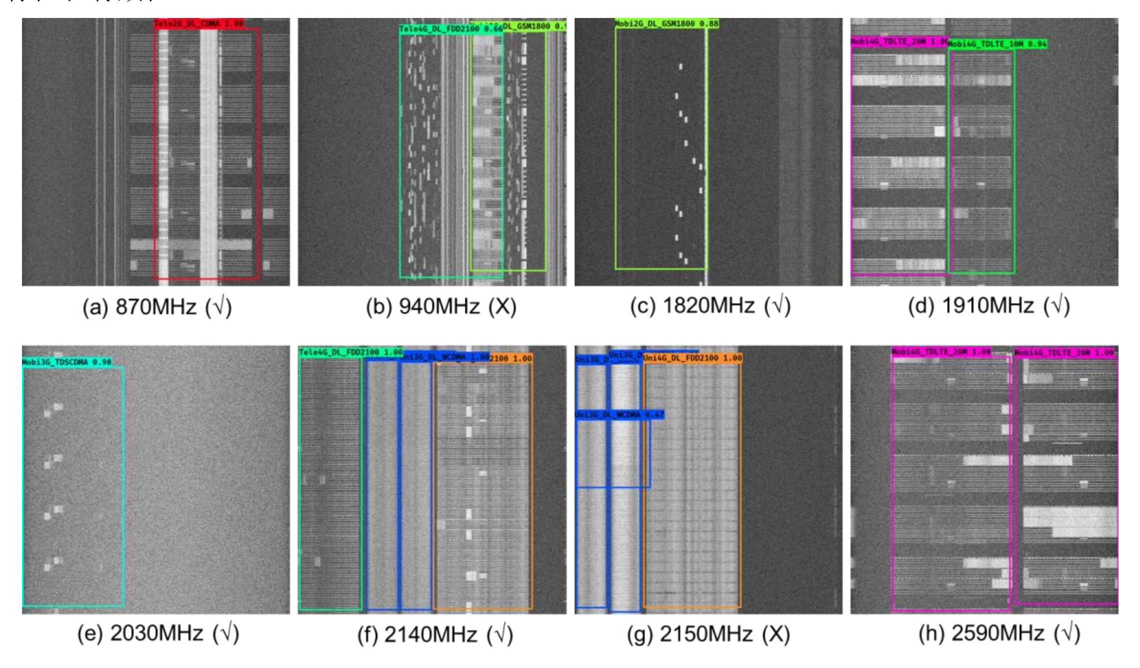


图 7 模型测试结果样本

为了更严谨的评估算法的性能，我们需要更多的量化指标。如图 8 所示，对于二分类问题，根据真实值和预测值的结果，可以将整个样本集分为 4 个象限，如图中灰色部分所示，分别是真正例（TP）、伪正例（FP）、真反例（TN）和伪反例（FN）。经过简单扩展，二分类混淆矩阵（confusion matrix）也可以变成多分类混淆矩阵。基于混淆矩阵，可以构造多种分类器性能评价指标。

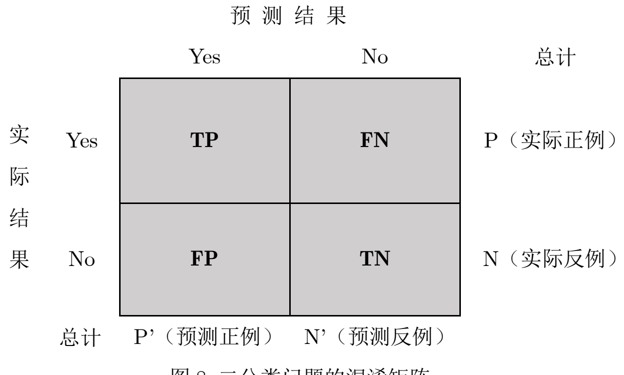


图 8 二分类问题的混淆矩阵

正确率（accuracy）是最常见的评价指标，其定义公式为 accuracy ＝ (TP+TN)/(P+N)，即被分对的样本数占总样本数的比例。通常来说，正确率越高，分类器越好。图 9 是本模型测试结果的混淆矩阵，其中纵轴表示真实标签，横轴表示预测标签，在 12×12 的坐标内，用方框颜色表示数值，颜色越深、数值越大。一般而言，混淆矩阵对角线的颜色越深越好，其它坐标点的颜色越浅越好。由图可见，模型对于大部分信号识别性能很好，但是容易将 Uni4G_DL_FDD1800 和 Uni4G_DL_FDD2100 分别识别成 Tele4G_DL_FDD2100 和 Tele4G_DL_FDD2100。根据混淆矩阵和定义公式，可以算出该模型的识别正确率为 86.04%。

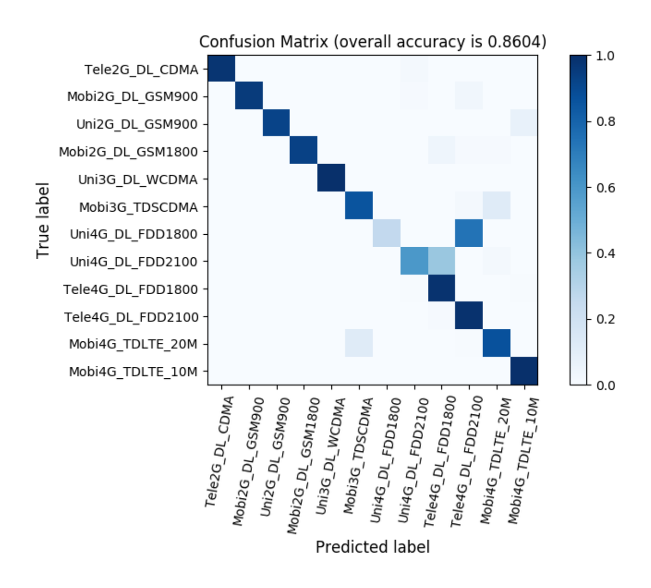


图 9 测试结果-12 类信号识别的混淆矩阵

正确率虽然是一个很直观的评价指标，但是有时候并不能完全反映一个算法的好坏，特别是在样本数据分布不均衡时尤为明显，因此我们又引入了功能更强大的 PR 曲线。PR 曲线，即查准率-查全率（Precision-Recall）曲线，其中纵坐标为 Precision，横坐标为 Recall。Precision 又称精准率，其定义公式为 precision=TP/(TP+FP)，表示被分为正例的示例中实际为正例的比例；Recall 又称召回率，其定义公式为 recall=TP/(TP+FN)，表示所有正例中被分对的比例。PR 曲线可以直观展示分类效果的好坏，在分类、检索等领域应用广泛，一般而言，曲线越靠近坐标（1，1），分类器性能越好。

但是，PR 曲线毕竟是一个向量，不方便进行数值比较，因此我们通常使用平均查准率(Average-Precision，AP)，即 PR 曲线下的面积进行分类效果评估。对于一个分类器而言，往往能够同时识别多个目标，而一条 PR 曲线及其 AP 值只能表示其中一种目标的分类效果，因此，可以再对所有 AP 值求平均，就得到了 mAP 值。目前，mAP 值已经成为检测算法的核心性能指标之一。

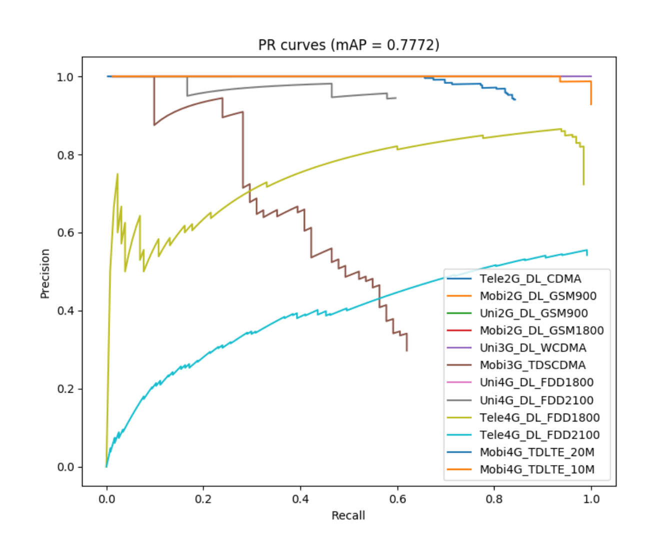


图 10 测试结果-12 类信号识别的 PR 曲线

图 10 是本模型测试结果的 PR 曲线图，表 5 是其对应的 AP 值。从图 10 可以看出，尽管 Uni4G_DL_FDD1800 和 Uni4G_DL_FDD2100 的识别精准度很高，均在 90%以上，但由于其召回率较低，分别约为 30%和 60%，因此整体的 AP 值指标并不太好，分别只有 26.02 和 58.50。换句话说，如果该分类器判定一个信号是 Uni4G_DL_FDD1800 或 Uni4G_DL_FDD2100，则很可能是对的，但也会出现很多漏检情况。Tele4G_DL_FDD2100 的情况刚好相反，其召回率可达到 100%，但精准度不高，只有不到 60%，也就是说，该分类器能够把所有的 Tele4G_DL_FDD2100 信号都检测出来，但也会出现很多误判情况。Mobi3G_TDSCDMA 的情况则兼具二者的特征，召回率和精准度都一般，整体识别效果也比较一般。除了这 4 种信号以外，其它 8 种信号的识别效果都比较好，某些信号的 AP 值甚至达到 95 以上。所有信号识别的平均 AP 值 mAP 达到 77.72，该指标非常接近目前在自然图像目标检测领域的最好指标——78.8[31]。

表 5 测试结果-AP 值

序号	信号名称	AP 值
1	Tele2G_DL_CDMA	97.67
2	Mobi2G_DL_GSM900	95.35
3	Uni2G_DL_GSM900	92.31
4	Mobi2G_DL_GSM1800	93.02
5	Uni3G_DL_WCDMA	100
6	Mobi3G_TDSCDMA	45.76
7	Uni4G_DL_FDD1800	26.02
8	Uni4G_DL_FDD2100	58.50
9	Tele4G_DL_FDD1800	85.09
10	Tele4G_DL_FDD2100	55.02
11	Mobi4G_TDLTE_20M	83.91
12	Mobi4G_TDLTE_10M	99.92
	mAP	**77.72**

6 结束语

无线电信号识别技术作为无线通信的一个重要研究领域，在军事电子战、政府频谱监管、商业频谱共享接入等方面应用广泛。结合其作为一种“模式识别”任务的本质，以及人工智能在模式识别任务中特有的优势，本文提出了一种基于图像深度学习的无线电信号识别算法——RadioImageDet 算法。该算法首先利用

传统数字信号处理方法，将 I/Q 采样数据流变换为一张二维频谱瀑布图，将无线电信号识别问题转化为图像识别领域的目标检测问题；进而利用人工智能在图像识别领域的先进成果，以 YOLOv2 模型为基础，构造了一种面向无线电信号识别的端到端深度神经网络模型——RadioYOLO 模型，并利用该模型识别信号的波形类型和时/频坐标。这种将无线电信号识别问题转化为图像目标检测问题的思路，在国内外尚属首次提出，给未来复杂电磁环境下智能化无线电信号识别技术的研究，开启了一条新思路。

为了检验算法的真实性能，本研究基于 GnuRadio 和 USRP 平台，开发了一套高速频谱采集系统，并在北京、天津等地实地采集频谱数据，构建无线电数据集。试验结果表明，RadioImageDet 算法能有效识别具象化后的无线电信号，识别效果与人眼近似，在实地采集的 12 种、4740 个样本的数据集中，识别准确度达到 86.03%，mAP 值达到 77.72，检测时间在中等配置的桌面计算机上仅需 33 毫秒，充分验证了思路的可行性和算法的有效性。

RadioImageDet 算法只是基于图像深度学习的无线电信号识别技术思路的第一次初步尝试，从试验结果中也能看出一些不如人意之处。未来，我们将进一步研究其它信号具象化方法和更加适合无线电信号特征的神经网络模型。

参 考 文 献

[1] Octavia A. Dobre. Signal Identification for Emerging Intelligent Radios: Classical Problems and New Challenges [J]. IEEE Instrumentation & Measurement Magazine, 2015, 18(2): 11-18.

[2] 姚亚峰, 黄载禄. 通信信号调制识别技术[J]. 通信技术, 2003, 6(1): 41-50.

[3] E. E. Azzouz, A. K. Nandi. Automatic identification of digital modulation types [J]. Signal Processing, 1995, 47(1): 55-69.

[4] J Lopatka, M Pedzisz. Automatic Modulation Classification Using Statistical Moments and a Fuzzy Classifier [C]. International Conference on Signal Processing Proceedings, 2000, 3(1): 1500-1506.

[5] A. K. Nandi, E. E. Azzouz. Algorithms for Automatic Modulation Recognition of Communication Signals [J]. IEEE Transactions on Communications, 1998, 46(4): 431-436.

[6] J. Reichert. Automatic Classification of Communication Signals Using Higher Order Statistics [C]. IEEE International Conference on Acoustics, Speech, and Signal Processing, 1992, 5:221-224.

[7] C. M. Spooner, A. N. Mody, J Chuang. Modulation Recognition Using Second- and Higher-Order Cyclostationarity [C]. IEEE International Symposium on Dynamic Spectrum Access Networks, 2017, 1: 1-3.

[8] S. S. Soliman, S. Z. Hsue. Signal Classification Using Statistical Moments [J]. IEEE Transactions on Communications, 1992, 40(5): 908-916.

[9] Y. Yamashita, H. Ochiai. A Classification of OFDM Signals With or Without DFT Precoding Based on High-Order Moment [C]. IEEE International Symposium on Dynamic Spectrum Access Networks, 2017, 1: 1-2.

[10] 陈健, 阔永红, 李建东, 等. 通信信号自动识别方法[J]. 电路与系统学报, 2005, 10(5): 102-109.

[11] D. L. Parnas. The Real Risks of Artificial Intelligence [J]. Communications of the ACM, 2017, 60(10):27-31.

[12] M. L. D. Wong, A. K. Nandi. Automatic Digital Modulation Recognition Using Spectral and Statistical Features with Multi-layer Perceptron [C]. Signal Processing and its Applications, 2001, 2: 390-393.

[13] G. Arulampalam, V. Ramakonar, et al. Classification of Digital Modulated Schemes Using Neural Networks [C]. International Symposium on Signal Processing and its Applications, 1999, 2: 649-652.

[14] K. Triantafyllakis, M. Surligas, G. Vardakis. Phasma: An Automatic Modulation Classification System Based on Random Forest [C]. IEEE International Symposium on Dynamic Spectrum Access Networks, 2017, 1: 1-3.

[15] T. O'Shea, J. Corgan, T. C. Clancy. Convolutional Radio Modulation Recognition Networks [C]. International Conference on Engineering Applications of Neural Networks, 2016, 1:213-226.

[16] T. O'Shea, J. Corgan, T. C. Clancy. Unsupervised Representation Learning of Structured Radio Communication Signals [C]. First International Workshop on Sensing, Processing and Learning for Intelligent Machines, 2016, 1:1-5.

[17] T. O'Shea, N. West, M. Vondal, et al. Semi-supervised Radio Signal Identification [C]. International Conference on Advanced Communication Technology, 2017, 1:33-38.

[18] J. L. Ziegler, R. T. Arn, W. Chambers. Modulation Recognition with GNU radio, Keras and HackRF [C]. IEEE International Symposium on Dynamic Spectrum Access Networks, 2017, 1:1-3.

[19] 于金华, 周羽丰, 陈柱文. Recent Trends of DARPA—RFMLS [Online]. http://m.sohu.com/a/192607899_466078. 2017.

[20] K. Karra, S. Kuzdeba, J. Petersen. Modulation Recognition Using Hierarchical Deep Neural Networks [C]. IEEE International Symposium on Dynamic Spectrum Access Networks, 2017, 1: 1-3.

[21] N. West, T. O'Shea. Deep Architectures for Modulation Recognition [C]. IEEE International Symposium on Dynamic Spectrum Access Networks, 2017, 1:1-6.

[22] N. West, K. Harwell, B. McCall. DFT Signal Detection and Channelization with a Deep Neural Network Modulation Classifier [C]. IEEE International Symposium on Dynamic Spectrum Access Networks, 2017, 1: 1-3.

[23] K. Arumugam, I. Kadampot, M. Tahmasbi, et al. Modulation Recognition Using Side Information and Hybrid Learning [C]. IEEE International Symposium on Dynamic Spectrum Access Networks, 2017, 1: 1-2.

[24] L. Huang, J. Yu, Z. Shen, et al. Radio Individual Identification Via Stable Communication Signals Based On Subordinate Component Analysis [C]. IET International Radar Conference, 2016, 1:1-4.

[25] Y. Yuan, Z. Huang, F. Wang, et al. Radio Specific Emitter Identification Based on Nonlinear Characteristics of Signal [C]. IEEE International Black Sea Conference on Communications and Networking, 2015, 1:77-81.

[26] K. Ahmad, G. Shresta, U. Meier, et al. Neuro-Fuzzy Signal Classifier (NFSC) for Standard Wireless Technologies [C]. International Symposium on Wireless Communication Systems, 2010, 1:616-620.

[27] M. Schmidt, D. Block, U. Meier. Wireless Interference Identification with Convolutional Neural Networks [J]. IEEE 15th International Conference on Industrial Informatics, 2017, 1:180-185.

[28] Zhang M, Diao M, Guo L. Convolutional Neural Networks for Automatic Cognitive Radio Waveform Recognition [J]. IEEE Access, 2017, 5(1):11074-11082.

[29] R. Girshick, J. Donahue, T. Darrell and J. Malik. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014, 1: 580-587.

[30] R. Girshick. Fast R-CNN [C]. IEEE International Conference on Computer Vision, 2015, 1:1440-1448.

[31] S. Ren, K. He, R. Girshick and J. Sun. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6):1137-1149.

[32] T. Lin, P. Dollar, R. Girshick，K. He, et al. Feature Pyramid Networks for Object Detection [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017, 1:936-944.

[33] J. Redmon and A. Farhadi. YOLO9000: Better, Faster, Stronger [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017, 1: 6517-6525.

[34] W. Liu, D. Anguelov, D. Erhan, et al. SSD: Single Shot Multibox Detector [C]. European Conference on Computer Vision, 2016, 1:21-37.

[35] B. Danev and S. Capkun. Transient-based Identification of Wireless Sensor Nodes [C]. International Conference on Information Processing in Sensor Networks, 2009, 1:25-36.

[36] C. Szegedy, W. Liu, Y. Jia, et al. Going Deeper with Convolutions [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2015, 1:1-9.

[37] K. He, X. Zhang, S. Ren, J. Sun. Deep Residual Learning for Image Recognition [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016, 1:770-778.

作者简介：



周鑫（1986-），男，河南项城人，博士，中科院软件所高级工程师、硕士生导师，主要研究方向为认知无线网络、无线信道接入、射频机器学习等。

何晓新（1966-），女，硕士，中科院软件所高级工程师，主要研究方向为认知无线电与信息系统。

郑昌文（1969-），男，博士，中科院软件所研究员、博士生导师，主要研究方向为智能信息处理、空间系统仿真。